

LOUDNESS OF HARMONIC
AND INHARMONIC TWO-TONE COMPLEXES

by

HOWARD LAWRENCE GOLUB

S.B., University of Virginia
1973

SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE
DEGREE OF MASTER OF
SCIENCE

at the

MASSACHUSETTS INSTITUTE OF
TECHNOLOGY

February, 1975

Signature of Author.....
Department of Aerospace Engr.

Certified by.....
Thesis Supervisor

Accepted by.....
Chairman, Departmental
Committee



ABSTRACT

LOUDNESS OF HARMONIC
AND INHARMONIC TWO-TONE COMPLEXES
by

Howard Lawrence Golub

"Submitted to the Department of Aeronautics and Astronautics on February, 1975 in partial fulfillment of the requirements for the degree of Master of Science."

Subjects were required to adjust eight comparison two-tone complexes to a two-tone standard. For the comparison sound containing two frequencies in a 2:3 ratio, the loudness was much softer than expected. It was also discovered that the loudness confusion (or standard deviation) resembled very closely the averaged equiloudness curves attained with the above paradigm. Loudness is a function of pitch for monotonies and this also seems to be true for two-tone complexes. Therefore, pitch confusion should be directly proportional to the loudness confusion which is directly proportional to the averaged equiloudness curves. Or, more simply, it was hypothesized here from the supporting evidence (without rigorous proof), that the loudness of two-tone complexes is a function of pitch ambiguity as well as intensity and pitch,

Thesis Supervisor: Dr.K.U.Ingard
Title: Professor of Physics

ACKNOWLEDGEMENTS

I would like to thank Dr. J.K. Haviland of the University of Virginia for his immeasurable influence in pointing me in the right direction, and Dr. A. Houtsma of the Massachusetts Institute of Technology for keeping me there.

TABLE OF CONTENTS

<u>Chapter No.</u>		<u>Page No.</u>
1	Introduction	5
2	Theory	10
3	Experimental Set-Up	16
4	Discussion of Data	24
5	Conclusion and Direction for Future Work	43
 <u>Appendix</u>		
A	Discussion of Combination Tones and Difference Tones	46
 <u>Figures</u>		
1	Focal Computer Program	17
2	Output Format	19
3	Curve of Zwicker, Flottorp, and Stevens	22
4, 5, 6, 7, 8	DIOT-DISS Case	25, 27-30
9, 10	DIOT-CONS Case	31, 32
11, 12, 13	DICH-DISS Case	34-36
14, 15	DICH-CONS Case	38, 39
16	Standard Deviation of Previously Reported Results	41
A	Amplitude and Phase of $2f_1-f_2$ CT	47
B	Amplitude and Phase of f_2-f_1 CT	48
<u>References</u>		49-51

Chapter 1

INTRODUCTION

An accepted definition of noise is: "a sound which lacks agreeable musical quality" or is noticeably unpleasant.¹ An excess of noise has been blamed for increased blood pressure, leading to heart attacks (first major killer in the United States), hyper-nervous tension, chronic anxiety, a decrease in sexual drive (which is serious in itself), and all around uneasiness. Although two-tone complexes (the stimuli in these experiments) are not the type of noise usually found in nature, it is a first step in quantitatively understanding exactly what noise really is.

The need to quantify the degree of noisiness or "musicality" of a sound becomes increasingly clearer, as the need to quantify annoyance becomes important. In order for one to legislate on the noise limits of any noise maker, we must first specify what is meant precisely by a noise level and how to measure it.

Because we are now living in this new era, A.E. ("After Einstein"), in order for us to define precisely what noise is, the obvious question that should first be answered is noise relative to who or what. This question can simply be answered psychoacoustically, as noise relative to one standard deviation (or approximately 70%) of the population. In other words, people decide what noise is quantitatively, and the answer to this question can only be achieved by focusing on perception. Although

this might seem to be an obvious point, it is not difficult to fall into the trap of thinking that a sound-level meter sets noise criteria instead of people.

Why is the noise from a jet aircraft so very disturbing and uncomfortable, while the music from a symphony by Beethoven is so beautiful and easy to listen to? The sound intensities (measured on A-scale sound level meters) at most parts of the symphony are very close to that of the aircraft, at an average height for landing maneuvers, yet one is considered as noise and the other as music. It was not an accident that the aforementioned definition of noise contained 'musical quality' to describe it.

There are really three reasons that can aptly explain this question. The first difference between music and noise is the time structure, where one sound has a more or less random time variation, and the other sound has a carefully defined time structure of which only a genius can construct. The second reason is the intensity variation, or dynamics. Again, the jet engine has a doppler type intensity variation at landing maneuvers, and of course, the music is carefully controlled. The last and most interesting difference will be called "spectral shape", or "frequency placements" within the sound. Although the jet noise has a large bandwidth, it should be noted that a symphony orchestra does also; it ranges from a piccolo to a bass fiddle, almost the entire audible range. It is readily seen that

the two spectra are obviously very different; while the music is comprised of the intricate summation of consonant (musically defined) sounds, the noise is the random summation of dissonant (again musically defined) sounds. It is this difference that this paper will attempt to focus on.

It is the writer's contention that there is a continuum between music and noise and that a point on this continuum can be defined by the three parameters of time, intensity, and frequency variations. Therefore, in order for us to define or quantify noise, music must inherently be defined or quantified also. Just exactly what do we mean by 'quantifying' music? Plomp² tried to do a comparison dealing with the percentage of chords where the rates were separated by a quarter critical bandwidth and a full critical bandwidth between a Mozart composition and a Schoenberg work. This was done to corroborate his ideas on consonance versus dissonance. Determinations were investigated as a function of critical bandwidth, but these methods still didn't really help to define music! Others have attempted to use methods or paradigms dealing with music to specify pitch, but these experiments just tend to imply things instead of defining them.³ There has been surprisingly little accomplished about the idea of music versus annoyance or loudness in any quantitative way. The only method feasible seems to be the same as that for noise,---the use of perception for quantification with respect to loudness or annoyance. The prevailing attitude of "You can't quantify an art form" seems to hinder progress. But again we must pose the question, how can we understand what

noise is without understanding its relation to music?

If "Led Zeppelin" had been heard twenty years ago, it would have been considered as distasteful and noisy to everyone; or if Mozart's music had been played by one of the ancient Romans, it would have certainly been considered as noise. People's tastes change with the changing times. You can even say that music is noise that has evolved over 5000 years! How then can music (an art form) be generally described and quantified? A closer look reveals that although the time and intensity variational structures have changed with each of these different modes of music, the spectral composition, or frequency placements per notes or unit sound (and for the most part the instruments) did not. The Western form of music (Oriental music has a different spectral structure which presents an interesting question as to physiological and psychological differences) always contained a special order of frequencies per note which is called tonality. Up until about sixty years ago, this was the only form really utilized by musicians and composers. With the advent of people like Schoenberg and Stravinsky, serialism was introduced which presents a different frequency structure. However, it should be noted that Stravinsky's most famous works (The Rites of Spring) were tonal, and his works containing serialism have yet to gain significant recognition, fifty years later.

It has been known since Pythagoras of ancient Greece that sounds containing harmonic frequency progressions are treated differently within the auditory system than are those

sounds of arbitrary frequency placements. People like J.L. Goldstein,⁴ Plomp,⁵ Houtsma,⁶ Schouten,⁷ and many others all have dealt with it in relation to pitch phenomena. It is now time to link the musicality of a sound to the subjective loudness or annoyance of that sound.

The difference between loudness and annoyance is indeed a subtle one, and in literature annoyance is usually associated with more complex sounds (found in nature). However, in these experiments they will be assumed the same. Through preliminary testing, where subjects were instructed to match loudness and/or annoyance, there were no statistically significant differences in the subject's response. This is not to say, however, that with other stimuli (more complex in nature) this difference would not be greater. Loudness will be used as the experimental parameter along with the aforementioned assumption.

Although the stimuli used in the following experiments are not natural in either presentation or frequency spectrum, the basic psychoacoustic assumptions should hold. These are in fact that subjects will respond to these stimuli as if they were natural and that the results may be generalized to apply to more complex stimuli with a little more proof.

The following pages contain the theory, experimental setup, data, discussion of data and direction for future work, which will serve to further clarify the effects of relative frequency placements within a discrete sound spectrum on subjective loudness.

Chapter 2

THEORY

The "Noisiness" (PNdB, Noys) of a monotone has been found by Kryter,⁸ to vary with frequency (Hz) at a constant selected intensity (measured in dB). The unit of loudness for single tones is called phons. For a more complex discrete sound, the methods of Loudness Summation (in sones) presented by S.S. Stevens⁹ or Zwicker's procedure for calculating loudness are as of now the major ways to assign a complex sound a particular loudness. Even these procedures are not regularly used; the A-scale, a poor representation of an inverted threshold curve used for its simplicity, is the law makers scale for most environmental noise such as traffic, street noise, etc. These methods handle the problem of the loudness of complex sounds from a standpoint of monotone, or narrow bandwidth loudness summations. Although spectral positioning is taken into account somewhat when making this summation, the relative positioning of the frequencies within the sound has been ignored. It should be quite evident that loudness summation shouldn't necessarily be a linear process. Phenomena such as combination tones, difference tones, and central nervous system effects all contribute to the probability that loudness summing is indeed a nonlinear process, of which simply adding the loudness of monotones does not yield a very accurate measure of subjective loudness.

For a sound containing more than one frequency, or narrow band of frequencies, the loudness might also be dependent on the structure or the ratios of these frequencies within the sound spectrum. Thus, a harmonic or consonant spectrum containing frequencies related by integer ratios could appear to be softer than a random or dissonant spectrum, even though the intensity of both sounds (A-scale, dBA) are the same.

A major reason for this idea could be explained by the combination tone and difference tone phenomenon. The combination tone, thanks to the work of J.L. Goldstein (1967)¹⁰ with his paper on "Auditory Nonlinearity", can be used to explain one possible reason for the difference in loudness between a dissonant and consonant sound. The combination tone appears at $2F_1 - F_2$, where F_1 and F_2 are the first and second frequencies of any ordered pair of frequencies in a discrete complex sound. For any harmonic sound spectrum of $f_1, f_2 \dots f_n$, where:

$$\begin{array}{ll} 1f_0 = f_1 & \\ 2f_0 = f_2 & f_0 = \text{fundamental} \\ \vdots & \\ nf_0 = f_n. & \end{array}$$

The combination tones as well as the difference tones $(F_2 - F_1)$, lie directly on a frequency lower than that pair in succession.

For example, if the sound were comprised of:

Sound..... f_1 — f_2 — f_3
 Combination tones..... $2f_1-f_2$ — $2f_2-f_3$
 Difference tones..... f_2-f_1 — f_3-f_2
 since $f_2 = 2f_1$, $f_3 = 3f_1$
 Overall sound..... f_1 — f_2 — f_3

In the special case of a sound containing two frequencies, as used in this experiment, of which they are in a 2:3 ratio, or the second and third partial of a harmonic series, the difference tone and the combination tone will have the frequency value of the fundamental of this series.

Although this seems to indicate that another frequency should be added somehow into the loudness summation, it was shown (again by Goldstein and J.L. Hall) that at the intensities used here (68 dB S.P.L.) and at the frequency separation (2:3), the combination tone as well as the difference tone would not be at all appreciable. There are also some complex interference phenomena¹¹ occurring between these two (depending on the relative phase of the primaries), since they are at the same frequency, which also tends to substantiate the above assumption.

In contrast however, for a dissonant sound of $f_1, f_2 \dots f_n$ where:

$$\begin{aligned} Af_0 &= f_1 \\ Bf_0 &= f_2 \\ &\vdots \\ Z_n f &= f_n \end{aligned}$$

$A, B \dots Z_n$ are, in general, not integers.
 f_0 = a common frequency

An example of this can be:

Sound..... f_1 — f_2 — f_3

Combination tones..... $2f_1-f_2$ — $2f_2-f_3$

Difference tones..... f_2-f_1 — f_3-f_2

since $F_2 = f_1 B/A$, $f_3 = f_1 C/A$.

Overall sound:

f_1 — $f_1(2-B/A)$ — $f_1(B/A-1)$ — $f_1 B$ — $f_1(2B/A-C/A)$ — $f_1(C/A-B/A)$ — f_1

It can be readily seen that the dissonant sound has more independent frequencies than the consonant, thereby creating more frequencies to sum into the overall loudness summation. It seems that the amplitude of the combination tones are proportional to the distance in Hz of F_1 and F_2 , as well as to the intensities, and the two primary frequencies would have to be fairly close together and with moderately high intensities for these combination tones to be heard.¹² The fact that a minor third (5:6) is just at the limits of producing audible combination tones seems to indicate that for most musical intervals, combination tones are usually not even audible, as is probably the case for the major fifth (2:3) used in these experiments.

In summing up, for a harmonic sound, the combination tones and difference tones if audible, always lie directly on one of the lower partials, while for a dissonant sound, they usually do not. The combination tones have amplitudes of 20-30 dB(depending

on primary frequency separation) below the main frequencies which indicate that when they lie directly on the main frequencies, they are masked and do not contribute significantly to the overall loudness. However, for a dissonant sound it would be very possible for the combination and difference tones to lie in areas (these areas can be identified by the single tone masking curves) where they would not be effectively masked and in which case they would have some effect on the overall loudness.

Although the calculation of loudness for complex sounds would probably be more correct when we include nonlinear aural harmonics, as we shall see from the following experiments, effects that arise in the central nervous system (or somewhere past the superior olivary complex) seem to be much more important. This finding should not seem too surprising due to the fact that pitch was found to be interpreted primarily in the central nervous system as shown in the work by A. Houtsma and J.L. Goldstein ("The Central Origin of the Pitch Complex Tones; Evidence from Musical Interval Recognition"). Although the above hypothesis utilizing combination and difference tones is very palatable and sounds as though it offers the answer, as we shall see, it is only a negligible part of the total effect.

We are now treading on that infinitely thin line between physiology and psychology when trying to explain an effect like this. It seems the more we learn (microscopically) about the central nervous system, the more we attribute specific perceptual

effects to "just" physiology, and that which we do not know very much about, such as the area past the neural joining of the two auditory systems (left and right ears), we attribute to psychology. Since this effect (consonant sounds appear softer than dissonant sounds) must occur in an area of the brain of which little is known microscopically, as we shall show later, we shall henceforth consider it a psychological effect, at least macroscopically speaking.

The fact that harmonically-oriented sounds do sound softer than dissonant sounds should not, of course, be assumed. The following experiments are devoted to the testing of this hypothesis. Although the following is not definite proof of this hypothesis, for nothing in perception is definite, the following experiments, we believe, offer a positive and convincing trend.

Chapter 3

EXPERIMENTAL SET-UP

The objective of these experiments is to find the relation between the loudness of a two tone complex and the relative position (in Hz) of their frequencies. For each test sound, the first frequency (f_1) was always set to a value of 1060 Hz, while the second (f_2), the parameter of the experiment, was variable. The intensities of both primary frequencies of the standard sound (1060-1108 Hz or 1060-1590 Hz) were always set to a constant 65 dB throughout the test.

With the aid of the PDP-12 computer and sound-proof room set-up (found at the Communications Bioengineering Group, R.L.E., at M.I.T.), the testing procedure was initiated as completely automated, where the computer controlled and tabulated the experiment (see Figure #1 for the actual focal computer program used). The experimental paradigm was designed to give the subject as much time as he or she needed to compare the loudness of a standard two frequency sound to a comparison two frequency sound. The subject was allowed to switch freely between the standard and the comparison sounds, heard diotically (condition in which the sound stimulus presented to each ear is identical through TDH-39 audiology-type earphones), at will. There were eight trials per test in which the standard was the

FIGURE #1

FOCAL COMPUTER PROGRAM

C-PSYCBL 1974 V-36

```

01.01 C ASK FOR NEC. INFO
01.05 T "SUBJECT NO.",!
01.10 A "WHAT KEYBOARD IS SUB.",N,X ANS(N,1)
01.11 S LB=0;S Q1=0;S P=8;S L1=0;S Z=0;S T=0;S W1=8
01.12 S U=90
01.15 A "WHICH OSCILLATES?",X,Y,!
01.16 A "WHICH ATTENUATORS?",A,B,!
01.17 A "WHICH SWITCH?",T1,!
01.18 X ATT(60,0,A);X ATT(60,0,B)
01.20 A "WHICH SOUND IS STAND. 1 FOR DISS.2 FOR CONS.?",S1,!
01.21 A "IS IT AREGULAR TEST 1-YES,2-NO?",K5,!
01.22 I (K5-2)2.01;
01.23 A "HOW MANY TEST FREQ. WHAT ARE THEY",W1,!
01.24 S P=W1
01.25 F J=1,1,W1; A B(J)

02.01 X ZER(0);X CLK(1000,0,0);X WAT(0);S Z=FANS((N-1)*4+1)-1
02.02 T Z1, Z,!
02.10 I (Z-8)2.01,2.3,2.01
02.30 T Z3,U,!
02.40 X LMP(0;0;8);X LMP(1;0;1)
02.50 I (S1-2)3.1,4.05;

03.10 C SET DISS AS STAND.
03.11 X ESW(0,0,T1);X ESW(1,1,T1)
03.14 X ATT(60,0,A);X ATT(60,0,B)
03.20 X OSC(1060,0,X);X OSC(1108,0,Y)
03.22 X ZER(0);X CLK(50;0;1)
03.25 X WAT(0)
03.30 X LMP(0,0,8);X LMP(1,0,1);GOTO 5.01

04.05 C SET CONS. S STAND.
04.07 X ESW(0,0,T1);X ESW(1,1,T1)
04.09 X ATT(60,0,A);X ATT(60,0,B)
04.10 X OSC(1060,0,X);X OSC(1590,0,Y)
04.12 X ZER(0);X CLK(50;0;1)
04.15 X WAT(0)
04.20 X LMP(0,0,8);X LMP(1,0,1);GOTO 5.01

05.01 I (K5-2)5.15,9.01;
05.15 S B(1)=1108;S B(2)=1225;S B(3)=1265;S B(4)=1333
05.20 S B(5)=1540;S B(6)=1590;S B(7)=1650;S B(8)=1800
05.30 X ZER(0);X CLK(100,0,0);X WAT(0);S Z=FANS((N-1)*4+1)-1
05.40 I (Z-3)5.5,5.6,5.5
05.50 I (Z-8)5.3,10.01;
05.60 I (LB-1)6.01,6.1,6.1

06.01 X LMP(0,0,8);X LMP(1,0,3)
06.02 I (T)6.05,6.05,6.1
06.15 S Q8=FRAN(P);S Q7=08
06.36 X ESW(0,0,T1);X ESW(1,1,T1)
06.10 X OSC(B(Q7),0,Y);X ATT(U+P+L1,0,A);X ATT(U+P+L1,0,B)
06.11 X ZER(0);X CLK(50;0;1)
06.12 X WAT(0)
06.15 X LMP(0,0,8);X LMP(1,0,3)
06.18 S T=T+1
06.20 X ZER(0);X CLK(100,0,0);X WAT(0);S Z=FANS((N-1)*4+1)-1
06.25 I (Z-14)6.3,7.02,7.02
06.30 I (Z-2)6.4,10.01;
06.40 I (Z-1)6.2,2.4,6.2

07.02 X ATT(P+L1+U,0,A);X ATT(P+L1+U,0,B)
07.03 S LB=LB+1
07.10 I (LB-1)7.2,7.3,7.2
07.20 X ZER(0);X CLK(100,0,0);X WAT(0);S Z=FANS((N-1)*4+1)-1

```

05.40 I (Z-3)5.5,5.6,5.5
 05.50 I (Z-8)5.3,10.01;
 05.60 I (L8-1)6.01,6.1,6.1

06.01 X LMP(0,0,8);X LMP(1,0,3)
 06.02 I (T)6.05,6.05,6.1
 06.15 S Q8=FRAN(P);S Q7=Q8
 06.06 X ESW(0,0,T1);X ESW(1,1,T1)
 06.10 X OSC(B(Q7),0,Y);X ATT(U+P+L1,0,A);X ATT(U+P+L1,0,B)
 06.11 X ZER(0);X CLK(50;0;1)
 06.12 X WAT(0)
 06.15 X LMP(0,0,8);X LMP(1,0,3)
 06.18 S T=T+1
 06.20 X ZER(0);X CLK(100,0,0);X WAT(0);S Z=FANS((N-1)*4+1)-1
 06.25 I (Z-14)6.3,7.02,7.02
 06.30 I (Z-2)6.4,10.01;
 06.40 I (Z-1)6.2,2.4,6.2

07.02 X ATT(P+L1+U,0,A);X ATT(P+L1+U,0,B)
 07.03 S L8=L8+1
 07.10 I (L8-1)7.2,7.3,7.2
 07.20 X ZER(0);X CLK(100,0,0);X WAT(0);S Z=FANS((N-1)*4+1)-1
 07.30 I (Z-14)7.6,7.5,7.4
 07.40 S L1=L1-2;GOTO 7.02
 07.50 S L1=L1+2;GOTO 7.02
 07.60 I (Z-8)7.7,10.01;
 07.70 I (Z-1)7.2,2.4,7.2

09.01 GOTO 5.3

10.01 S Q1=Q1+1;S L3=L3+1;S R=0;S I=0
 10.02 X ESW(0,0,T1)
 10.05 X LMP(0,0,8);X LMP(1,0,2)
 10.10 S L=L1+P+U-60
 10.11 S L1=0
 10.12 I (Q1-2)10.14,10.15,10.16
 10.14 S U=90;GOTO 10.2
 10.15 S U=50;GOTO 10.2
 10.16 I (Q1-4)10.17,10.18,10.20
 10.17 S U=40;GOTO 10.2
 10.18 S U=35;GOTO 10.2
 10.20 I X3.02, "FOR TEST",Q1,"WITH A SECOND FREQ. OF" 24,B(Q7),I
 10.25 T X3.02,"F(2)-F(1) =",B(Q7)-1060,I
 10.30 T X3.02, "ATTENUATION WAS",.25*L,"DB",I
 10.31 I (Q1-4)10.38,10.38;
 10.32 I (Q1-6)10.33,10.34,10.35
 10.33 S U=30;GOTO 10.38
 10.34 S U=30;GOTO 10.38
 10.35 S U=90
 10.38 I (Q1-W1)10.39,12.01;
 10.39 S C(Q1)=Q7
 10.40 I (Q7-P)10.5,10.8;
 10.41 S R=0
 10.51 S Q8=FRAN(P)
 10.52 S R=R+1
 10.53 I (C(R)-Q8)10.55,10.41,10.55
 10.55 I (R-Q1)10.51,10.61
 10.60 C THIS IS FOR NEW SOUND ASSIGNMENT
 10.70 GOTO 10.9
 10.80 S P=P-1
 10.82 S R=0
 10.85 S Q8=FRAN(P)
 10.86 S R=R+1
 10.87 I (C(R)-Q8)10.88,10.82,10.88
 10.88 I (R-Q1)10.86,10.91
 10.90 S Q7=Q8
 10.91 T X3,U,I
 10.92 I (Q1-W1)2.4,12.01;

12.01 X LMP(0,0,8);X LMP(1,0,4)
 12.10 QUIT

*

same sound throughout. When the subject felt that he or she had adequately equalized the 'loudness' of the comparison sound to that of the standard, he or she would then depress a trial-ending button, which automatically selected a new comparison sound with a new initial intensity. The subject was instructed to continue the same procedure with the new comparison sound as he did with the previous one until all eight comparison sounds were heard and analyzed.

For each trial the initial intensity of the comparison sound was randomized along with the order of presentation of each of the eight stimuli per test. The attenuators were adjusted to step .5dB per press, or .5dB/second if the attenuator button was held down. There were basically two intensity controls or attenuator buttons,---one for positive and one for negative attenuation. When the trial ending button was pressed, the computer automatically took the total adjusted attenuation of the comparison and subtracted from it the constant attenuation of the standard and then generated this number (Intensity $\frac{\text{comp. sound}}{\text{stand. sound}}$). The standard has a constant S.P.L. of 68 dB S.P.L. An example of the output format for one test can be seen in Figure #2.

The test computer program has the ability of presenting any of the two standards per test. The first, which shall be named as the "dissonant stand.", had a second frequency of 1108 Hz,---exactly one fourth of a critical bandwidth (the size of the critical bandwidths measured by Zwicker, Flottorp, and

FIGURE #2

SUBJECT NO. G.M.

WHAT KEYBOARD IS SUB.:1

WHICH OSCILLATES?:1 :2

WHICH ATTENUATORS?:1 :2

WHICH SWITCH?:1

WHICH SOUND IS STAND. 1 FOR DISS, 2 FOR CONS.:2

IS IT A REGULAR TEST 1-YES, 2-NO?:1

-1

-1

-1

-1

-1

-1

8

80

FOR TEST 1.00 WITH A SECOND FREQ. OF 1265

$F(2)-F(1) = 205$

ATTENUATION WAS 1.00DB

90

FOR TEST 2.00 WITH A SECOND FREQ. OF 1800

$F(2)-F(1) = 740$

ATTENUATION WAS 4.00DB

50

FOR TEST 3.00 WITH A SECOND FREQ. OF 1333

$F(2)-F(1) = 273$

ATTENUATION WAS 0.25DB

40

FOR TEST 4.00 WITH A SECOND FREQ. OF 1590

$F(2)-F(1) = 530$

ATTENUATION WAS 0.25DB

85

FOR TEST 5.00 WITH A SECOND FREQ. OF 1650

$F(2)-F(1) = 590$

ATTENUATION WAS 1.00DB

30

FOR TEST 6.00 WITH A SECOND FREQ. OF 1060

$F(2)-F(1) = 0.00$

ATTENUATION WAS 0.00DB

80

FOR TEST 7.00 WITH A SECOND FREQ. OF 1540

$F(2)-F(1) = 480$

ATTENUATION WAS 2.50DB

90

FOR TEST 8.00 WITH A SECOND FREQ. OF 1225

$F(2)-F(1) = 165$

ATTENUATION WAS 1.00DB

Stevens in 1957)¹³ away from 1060 Hz, the first frequency. The reason for this selection of a second frequency was taken from the work by Plomp and G.F. Smoorenberg¹⁴ in their definition of maximum dissonance as a function of critical bands. The second, or "consonant standard", had a second frequency of 1590 Hz or exactly the third partial of a fundamental of 530 Hz with 1060 Hz (or f_1) being the second partial. Or, the two frequencies were in a ratio of 2:3 and a frequency distance larger than a critical bandwidth of approximately 180-190 Hz at this frequency range, also in agreement with Plomp's specifications.¹⁵

The second frequency of the comparison sounds were chosen strategically to approximate the same trends achieved by Zwicker and Stevens (1957)¹⁶ in their work on loudness summation and its relation to critical bands. In effect, their results showed that if a complex sound (in their experiments, four frequency complexes were used with the frequency difference of the first and last frequency defined as the bandwidth) had a bandwidth less than that of a critical band, it would be equal in loudness to that of another sound with a bandwidth or ΔF within that same critical band. As the ΔF of the complex sound increased, the subjective loudness also increased monotonically, at least up until the largest ΔF tested or about 1500 Hz. It should be noted that in this paper Stevens did make some reference to the fact that some spectra he tested, when they contained equal intervals, seemed to be perceived as louder. Although at first

glance this seems to be in contradiction with the ideas presented here, however, from a closer inspection of the frequency values, it can be seen that these spectra were not harmonic progressions; they just had equal intervals, and were not in integer ratios.

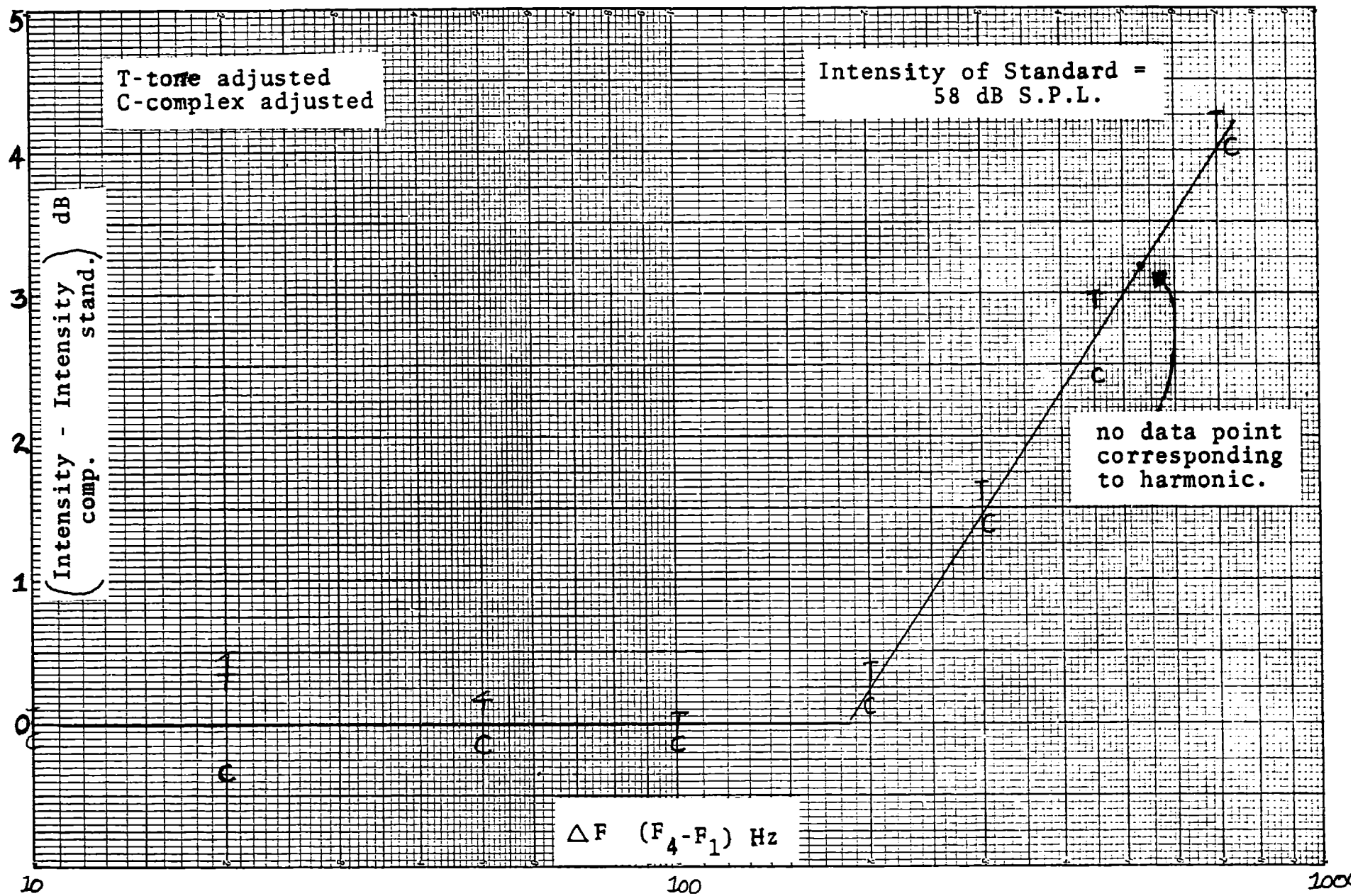
The curve of loudness versus ΔF (refer to Figure #3) that they had attained was somewhat dependent on intensity of their standard. However, at 65-70dB, the curve was most well behaved and pronounced.

There are essentially three differences between this experiment and ours. First, the standard they had used was a pure tone (1000 Hz), so the subject had to compare a complex sound to a pure tone. This can be considered a mild form of cross-modality testing which can only add confusion. We are not affected by pure tones as we are complex sounds; if we were, the whole problem would indeed be quite simple.

The second major difference is our concentration on ΔF 's corresponding to harmonic or consonant progressions.¹⁷ If the ideas presented before are correct, then a very narrow (frequency-wise) dip about 5-10 Hz wide, should be seen in the curve at a ΔF of a harmonic. Indeed, if one is not looking specifically for it, it could very easily be overlooked, and if seen once or twice could be dismissed as experimental error. The reasons for the dip being so narrow at the consonances can possibly be explained by the concept of mistuned consonances, where two frequencies very close to a harmonic

Figure #3.

Curve of Zwicker, Flotterp, and Stevens (1957)



ratio could produce beating and therefore sound dissonant .¹⁸ For example, if a first frequency was set to 1060 Hz and the second was mistuned to 1598 Hz (mistuned by 8 Hz), this would not necessarily sound as pleasing or musically consonant.

The third and final difference between Steven's experiment and ours, is that he used four frequencies in his comparison sound to our two. This factor does not seem to be very important due to some preliminary tests in which we presented to the subjects through the same experimental paradigm, two four tone stimuli,---the first being a consonant progression (500-1000-1500-2000 Hz) with a ΔF of 1500 Hz, and the second a dissonant stimulus (500-633-1745-2000 Hz) also with a ΔF of 1500 Hz. The subjects consistently adjusted the dissonant stimuli softer (or heard them as louder) by 4-5dB. This result seems to indicate that the effect gets progressively larger with the increasing number of component frequencies up until at least four (we did not test for more than this), and that if we had used four frequencies in these experiments, the dip would probably have been even more pronounced. It should also be noted that the loudness of two, three tone complexes were compared (in the Senior Thesis, "A Look at Noise and its Effect on Man", by H.L. Golub), and the loudness differential was approximately 2-3 dB.

The following data does show the monotonic upward trend as that achieved by Stevens and Co., however, with one interesting difference,---the area around the ΔF corresponding to the harmonic.

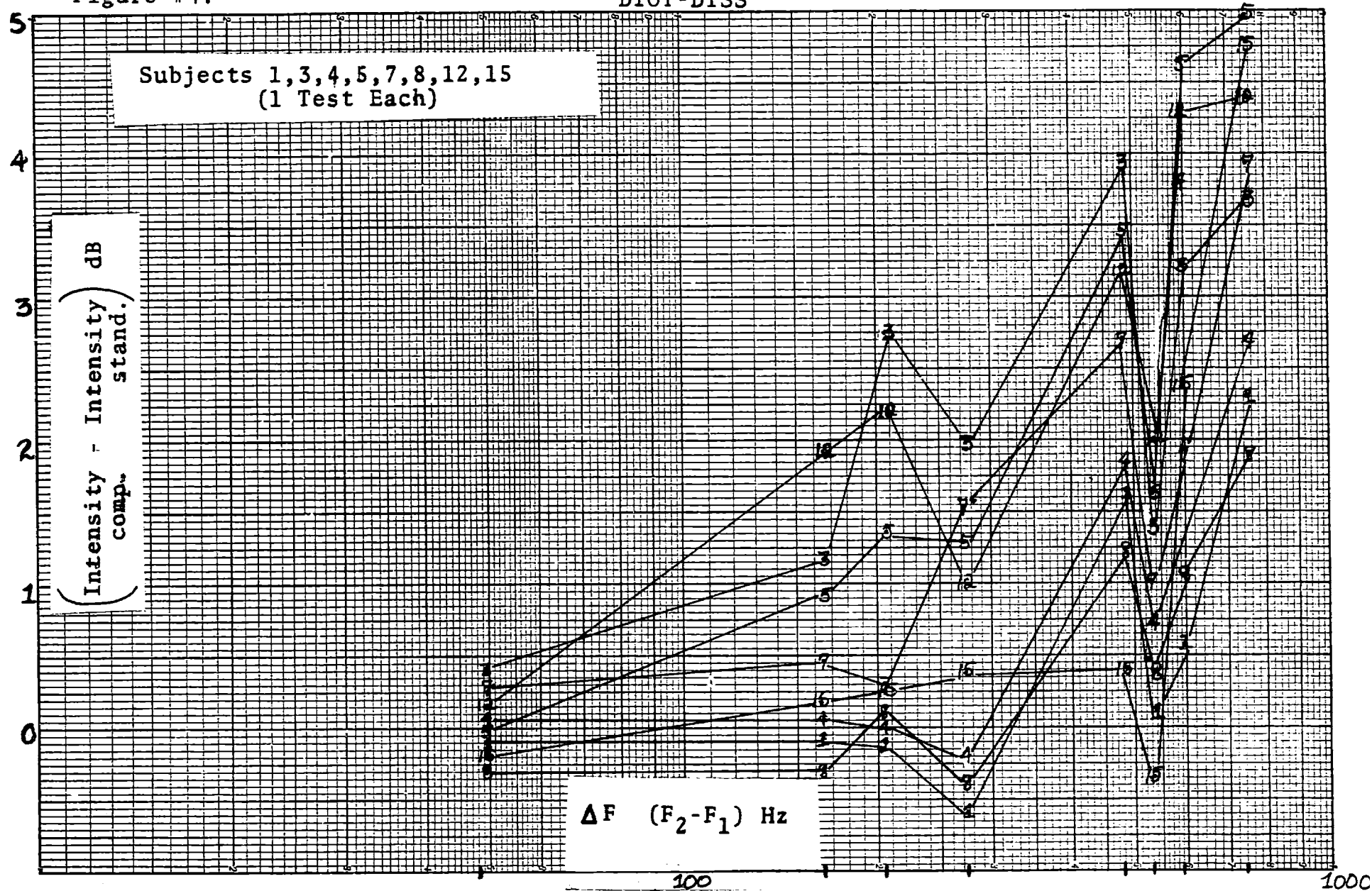
Chapter 4

DISCUSSION OF DATA

The eight comparison sounds have second frequencies (F_2) of 1108, 1225, 1265, 1333, 1540, 1590, 1650, 1800 Hz, and a first frequency (F_1) of a constant 1060 Hz. Hence, for the dissonant standard, a comparison sound with a F_2 equal to 1108 Hz was the control of the experiment, and for the consonant standard, an F_2 equal to 1590 Hz was the control. The first group of tests involved the case where all the stimuli were introduced diotically, with half consisting of the dissonant standard and the other half, the consonant standard. Figure #4 contains the results obtained from eight subjects, where the $\left(\begin{matrix} \text{Intensity} & - & \text{Intensity} \\ \text{comp.} & & \text{stand.} \end{matrix} \right)$ dB is the ordinate and the $\log_{10}(\Delta F)$ Hz, or the \log_{10} of the frequency difference of $F_2 - 1060$ Hz as the abscissa. All of these curves are the result of diotically introduced stimuli with a dissonant standard. It is readily seen that there is indeed a 'dip' at a ΔF corresponding to a $F_2:F_1$ of 3:2 for each of the subjects tested. There is also the upward trend noted in Steven's work after the critical band, corresponding to approximately 160-200 Hz. Although the slope of these upward trends vary, this is to be expected due to the fact that these curves were obtained after only one test for each subject. These representative curves are presented to show the repetition and consistency of this 'dip', and that from subject to subject, it appears quite visably.

Figure #4.

DIOT-DISS



Altogether twenty-two subjects were tested with varied *number* of tests per subject. There were eighteen subjects who were just tested once (Refer to Figure #4 for eight of these), and out of the curves of those eighteen, sixteen showed definite 'dip' type patterns at the consonant, while the other two were ambiguous. Out of the remaining four, each was tested twenty, eighteen, eighteen and ten times respectively. Out of these, a total of sixty-six, only three were without these 'dips' at the harmonic for the diotic case. In summing up, for twenty-two subjects of which there were a total eighty-four of these curves, only five did not effectively show the 'dip', or 94.1% of all those tested effectively did.¹⁹ Curves showing the average of those four tested more than once for the diotic dissonance (diot-diss) case can be seen on Figures #5-#8. It can be readily seen that the 'dip' is most apparent here.

The other half of this first group of tests consisted of the diotic case with a consonant standard (diot-cons). The curves for two of the four subjects for this case can be seen on Figures #9 and #10. The curves are representative of all those tested and there doesn't seem to be any contradiction between this case and that of the dissonant standard. The overall levels of the 'dip' are lower for the consonant standard case, but the trend is still very apparent.

As of now, there have been no gross deviations between

Figure #5.

DIOT-DISS

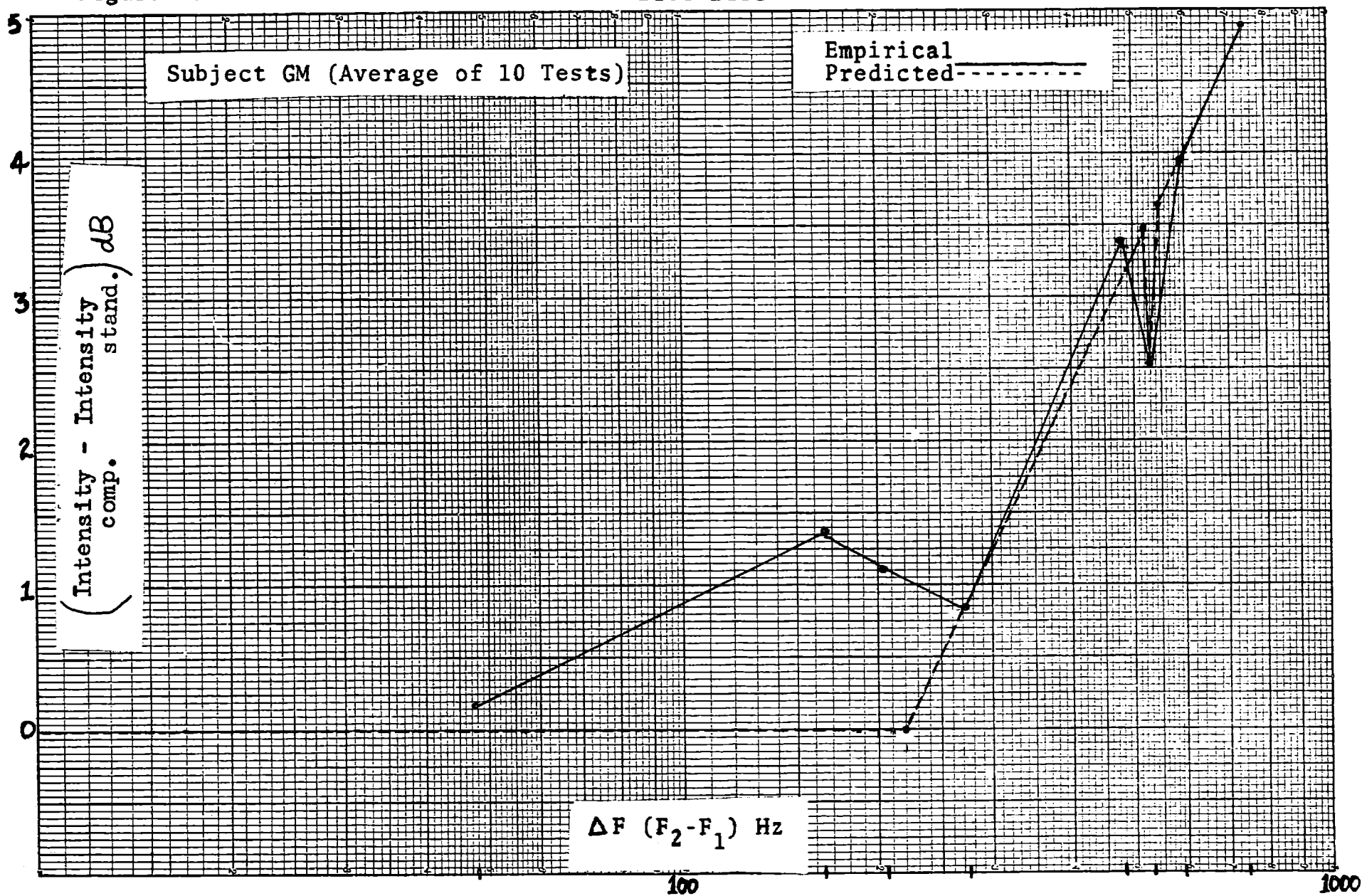


Figure #6.

DIOT-DISS

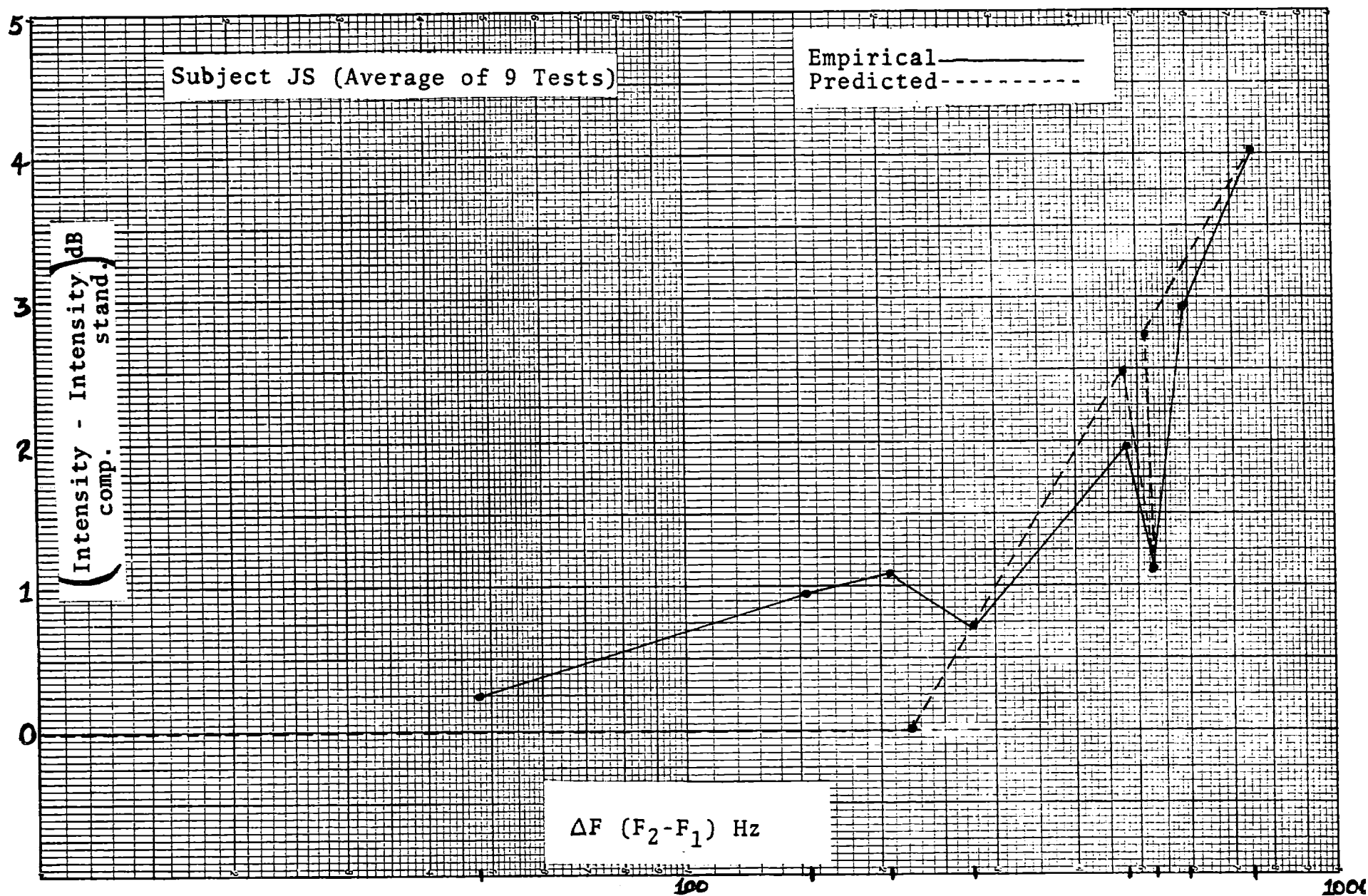


Figure #7.

DIOT-DISS

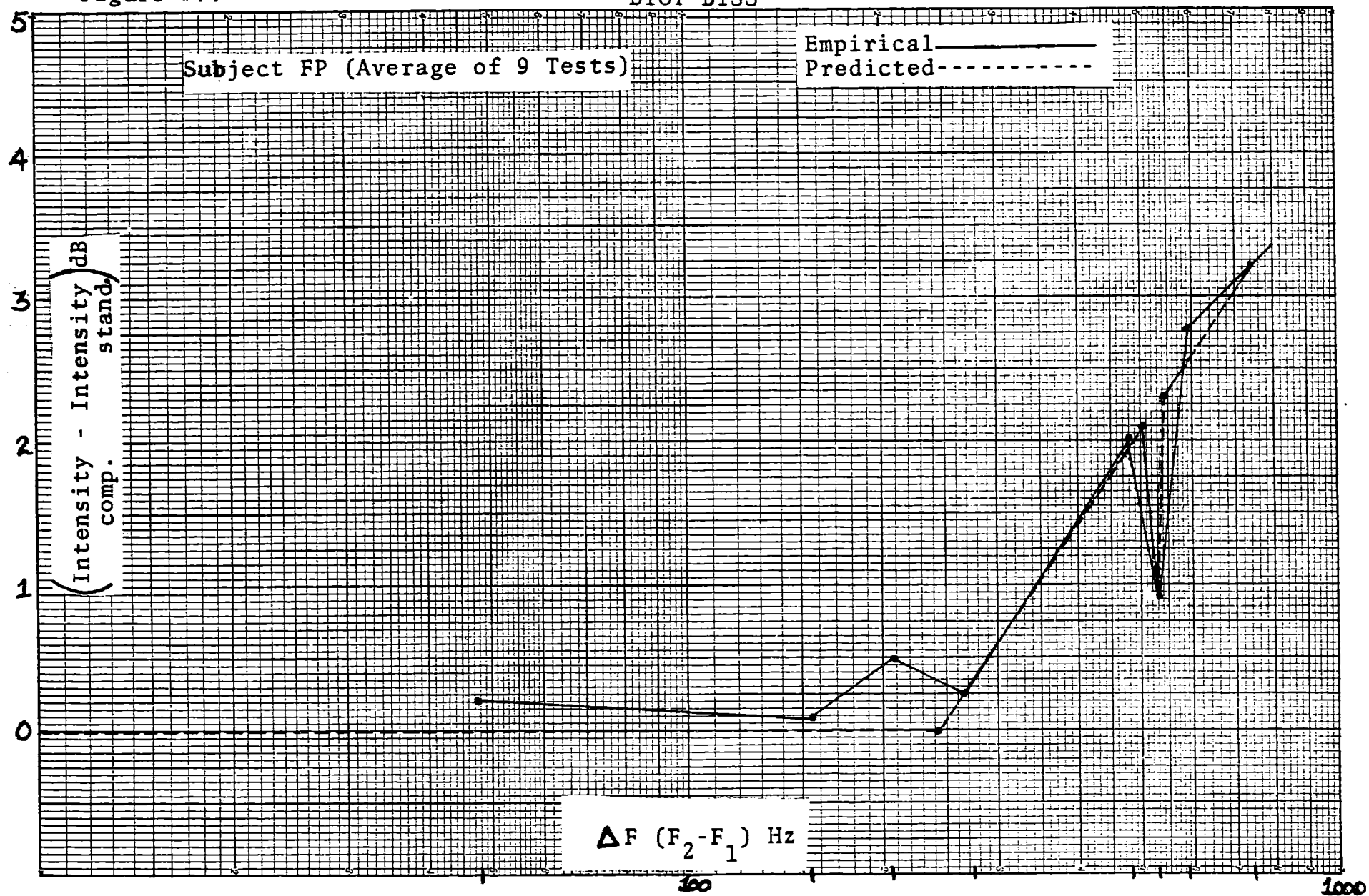


Figure #8.

DIOT-DISS

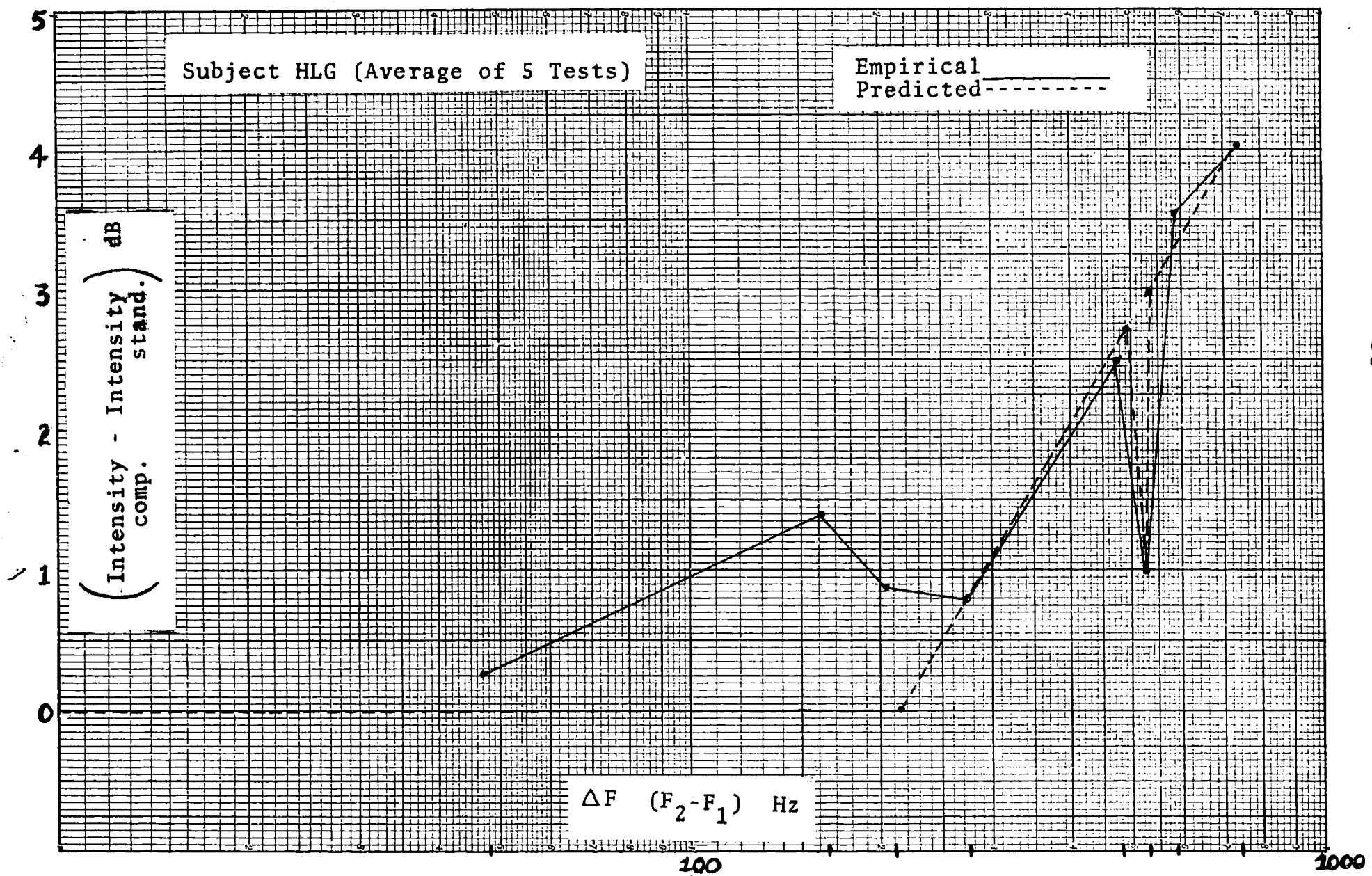


Figure #9.

DIOT-CONS

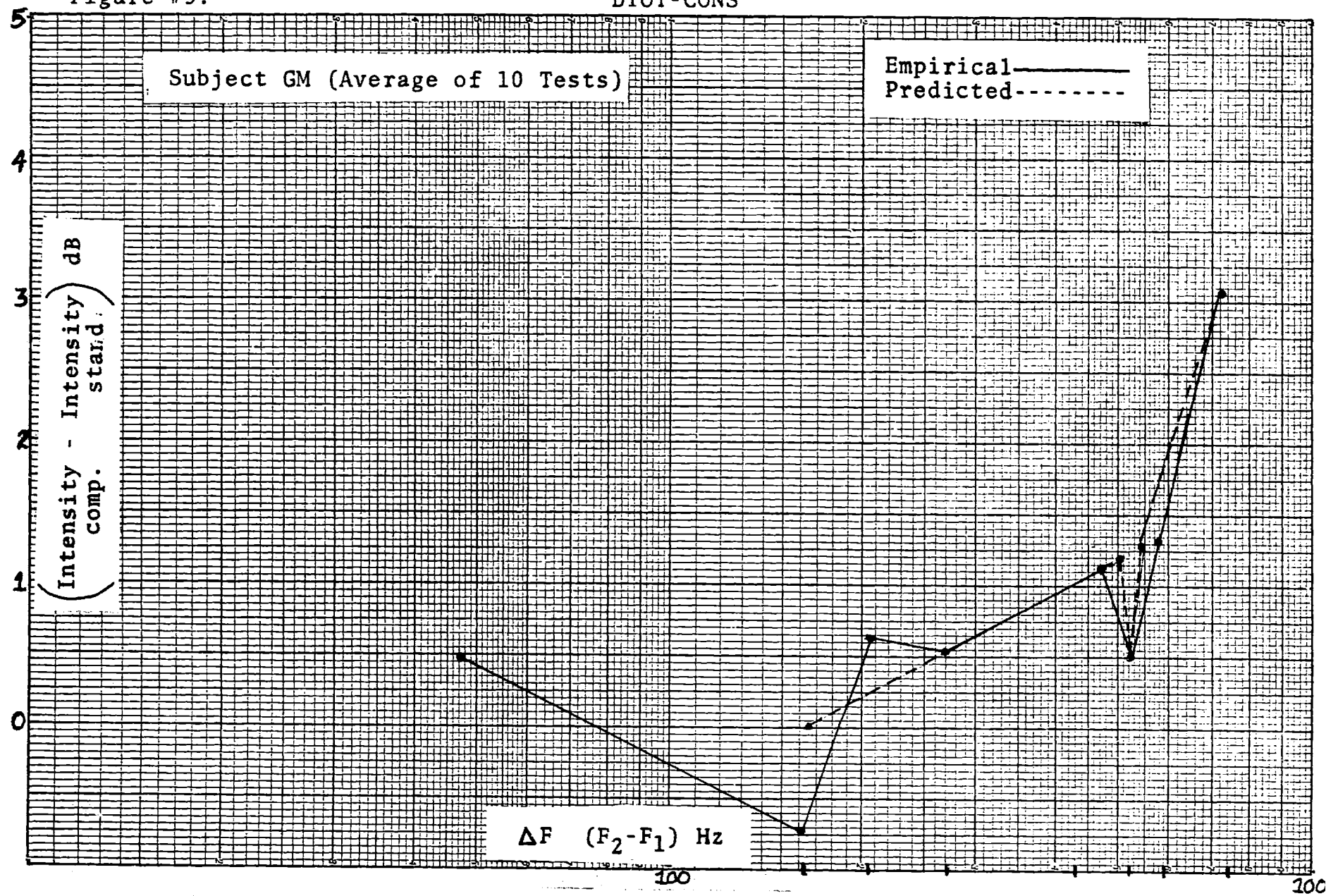
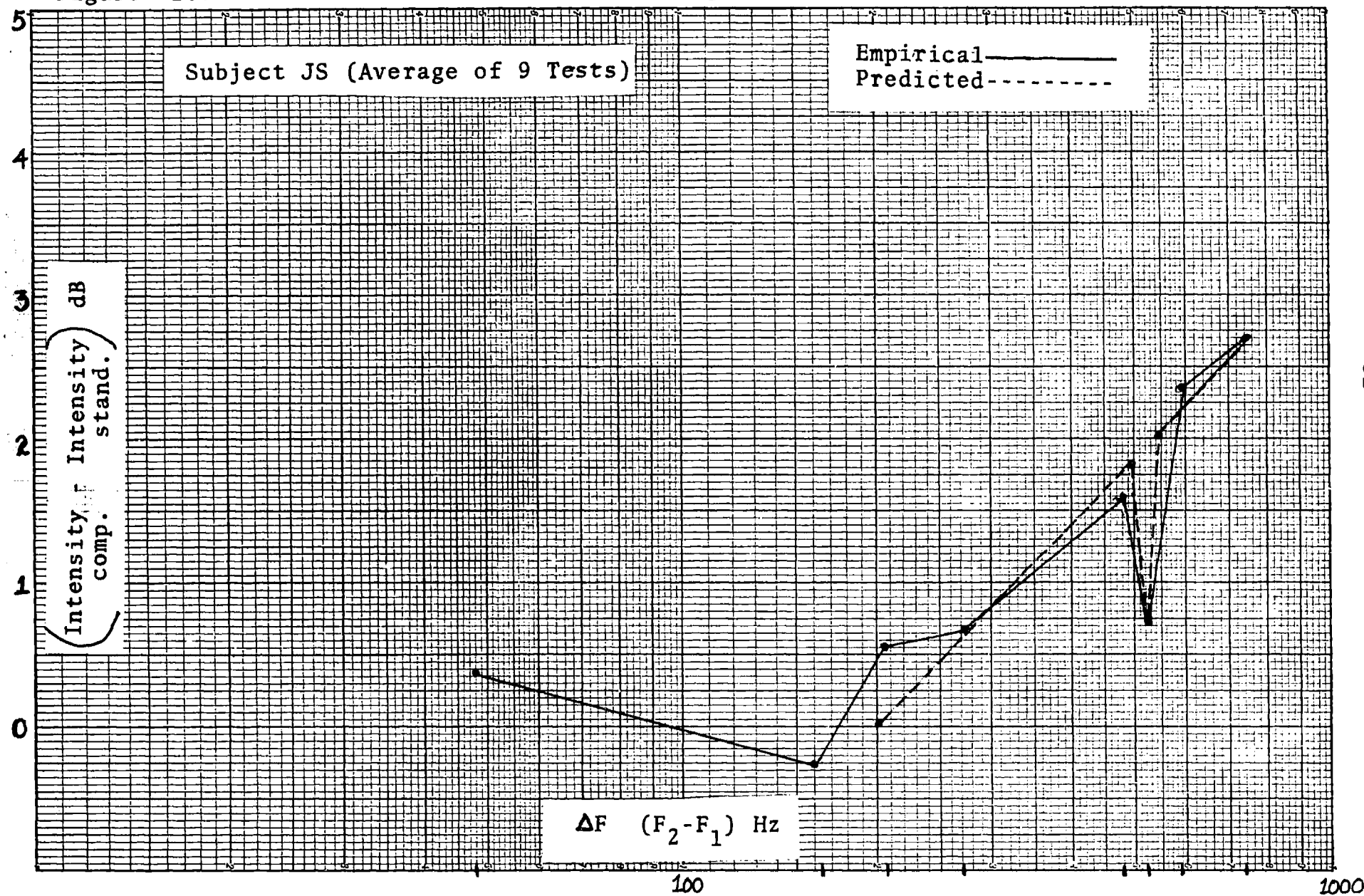


Figure #10.

DIOT-CONS



the empirical and predicted results. It was concluded that the combination tone hypothesis for the explanation of the 'dip' would be tested by introducing the stimuli dichotically (each frequency to a different ear). Since it is known that combination and difference tones are a monaural effect,---or more specifically, originating at the basilar membrane of the inner ear, if there is only one frequency introduced to each ear at a time, there would be no combination tones present.²⁰ The pitch of the sound would be the same; however, the confusion would vary in some complicated manner. This shall be discussed further later on in this paper.

For all the following tests, the standard sounds as well as the comparison sounds were all introduced dichotically. There were three subjects tested ten, nine, and nine times each for the dissonant standard and two subjects tested five times each for the consonant standard. Figures #11-13 show the curves for the dichotically introduced stimuli with a dissonant standard (dich-diss). The curves are most striking with respect to the 'dip' in that they are not affected to any significant degree by the fact that there are no nonlinear tones present. If anything, the 'dip' is even more pronounced. This fact seems to imply quite strongly that this harmonic effect seems to be the result of some central nervous system phenomena, or at least somewhere past the neural joining of the two ears (superior olivary complex). It probably occurs more centrally than this due to some psychoacoustical and neurophysiological reasons which are beyond the scope of this paper. At any rate,

Figure #11.

DICH-DISS

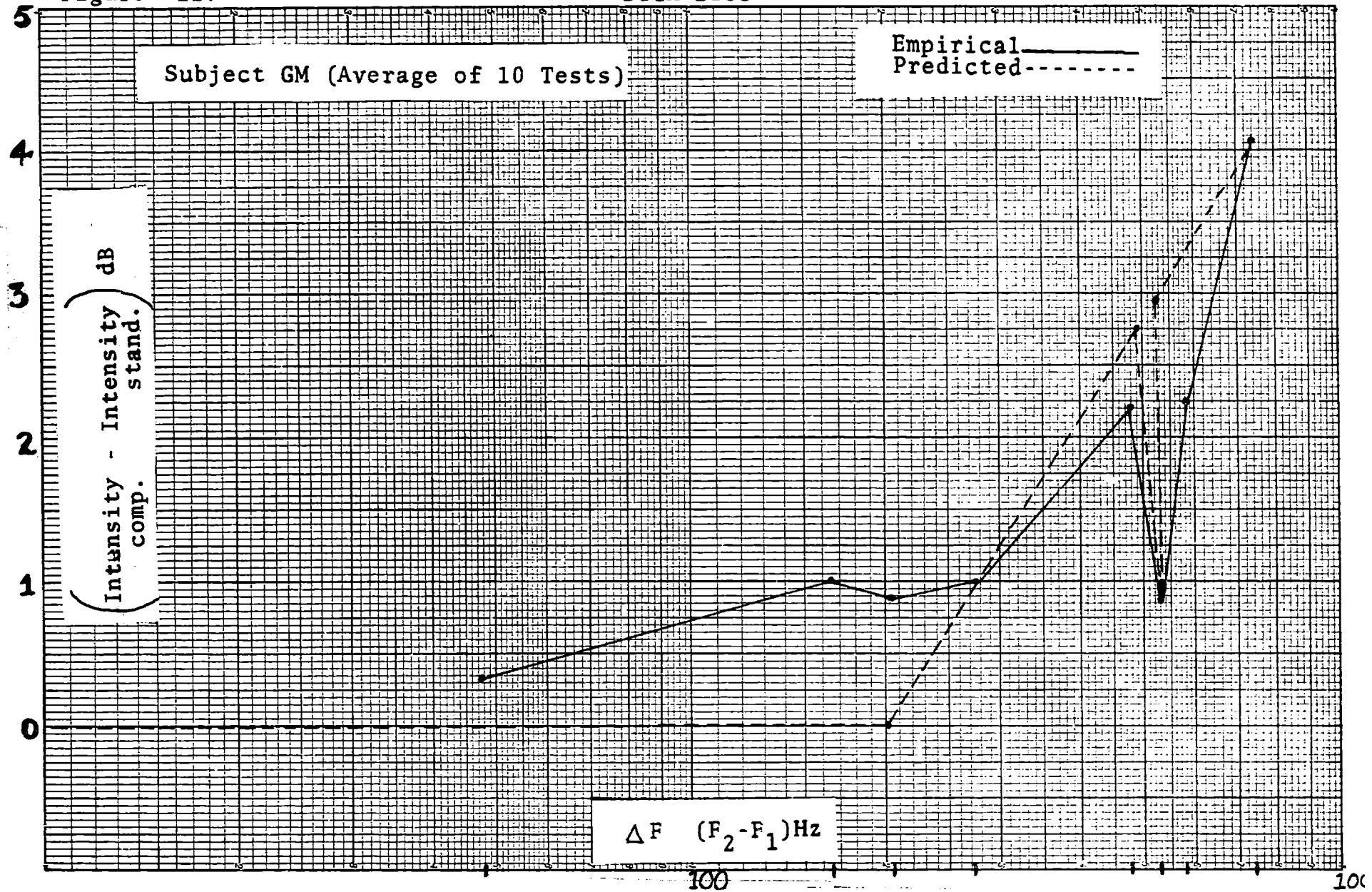
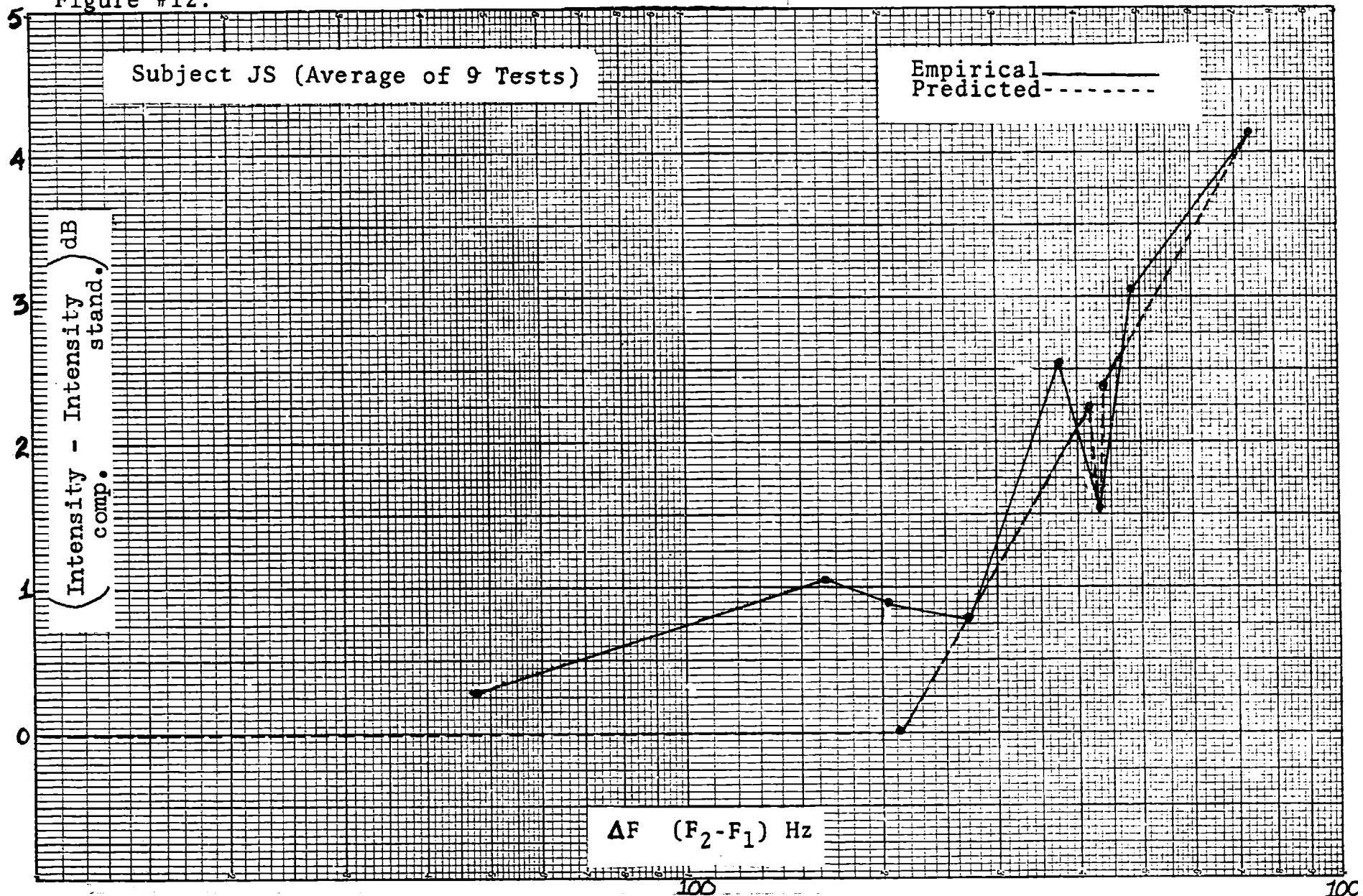


Figure #12.

DICH-DISS

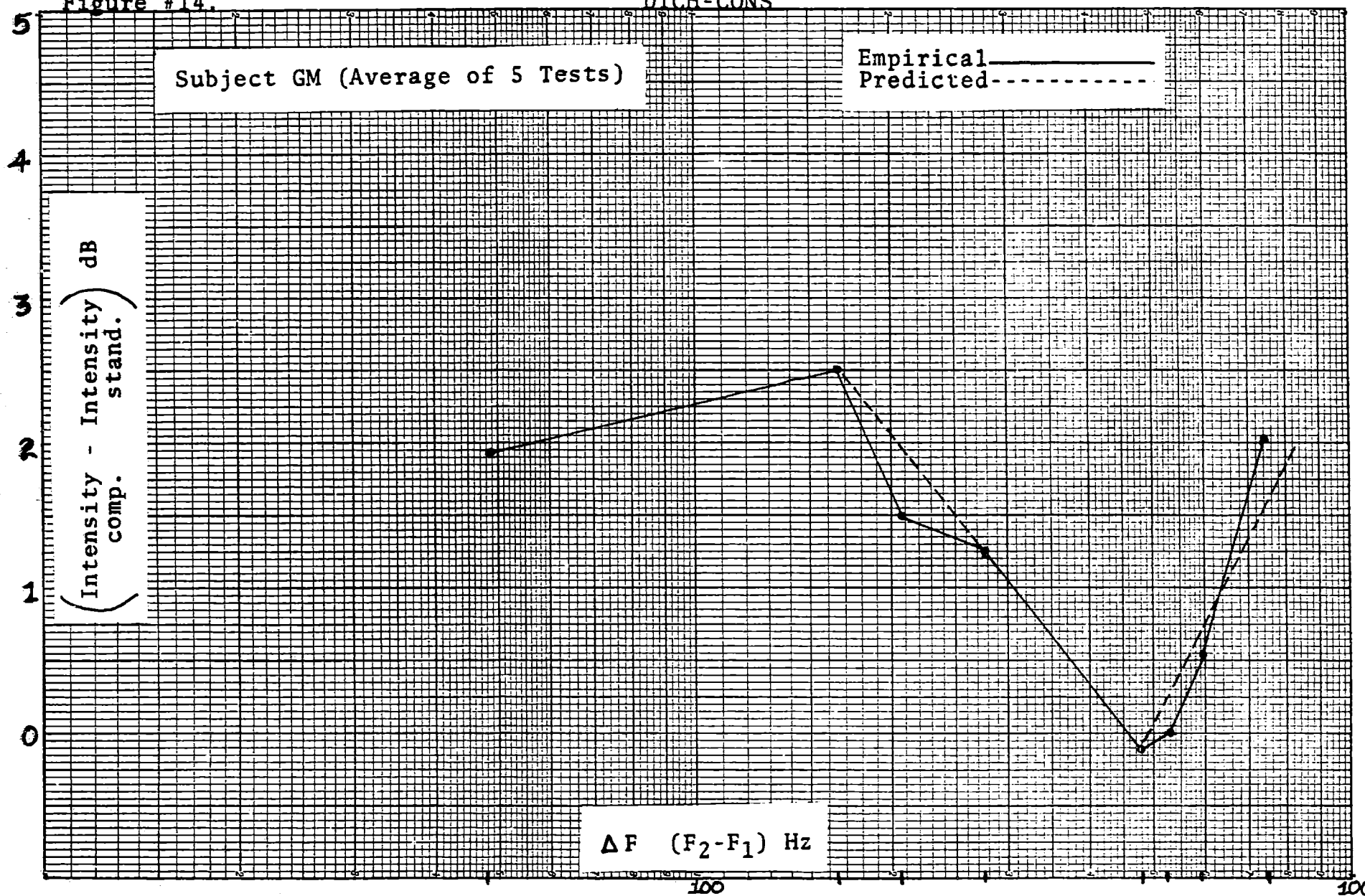


as stated earlier, these dichotic tests indicate that most of the effect must be attributed to an area of the brain about which little is known and as a matter of definition, it is therefore a psychological effect.

The curves shown on Figure's #14 and #15 are those for the dichotic presentation of the stimuli with a consonant standard. They are the most ambiguous of all four cases and indeed are not easily explained. It is these curves however, that prove to be the most interesting. These curves seem to indicate that although the value of the intensity - intensity comp. stand. was lower at the harmonic than that at the dissonant, there was no apparent 'dip' at the consonant. Instead there is a very large depression centered on the ΔF corresponding to the consonant. The possible explanation of this fact and the fact that for the diot-cons case, the dissonant end of the curves was not elevated above the value of the consonant, as is here, lead us to conclude that we are not using the correct parameters. If instead of using the ΔF for the abscissa, we used the psychoacoustic value of pitch and the relative confusion (or more simply, the difference between the standard deviation of each pitch determination of the standard subtracted from the comparison) as the parameters, these equiloudness curves should be more consistent under any case. The idea that the loudness of a discrete complex sound is a function only of the intensity, pitch, and pitch confusion is of course, a new one and has not been directly proven here. However, the standard deviations of

Figure #14.

DICH-CONS

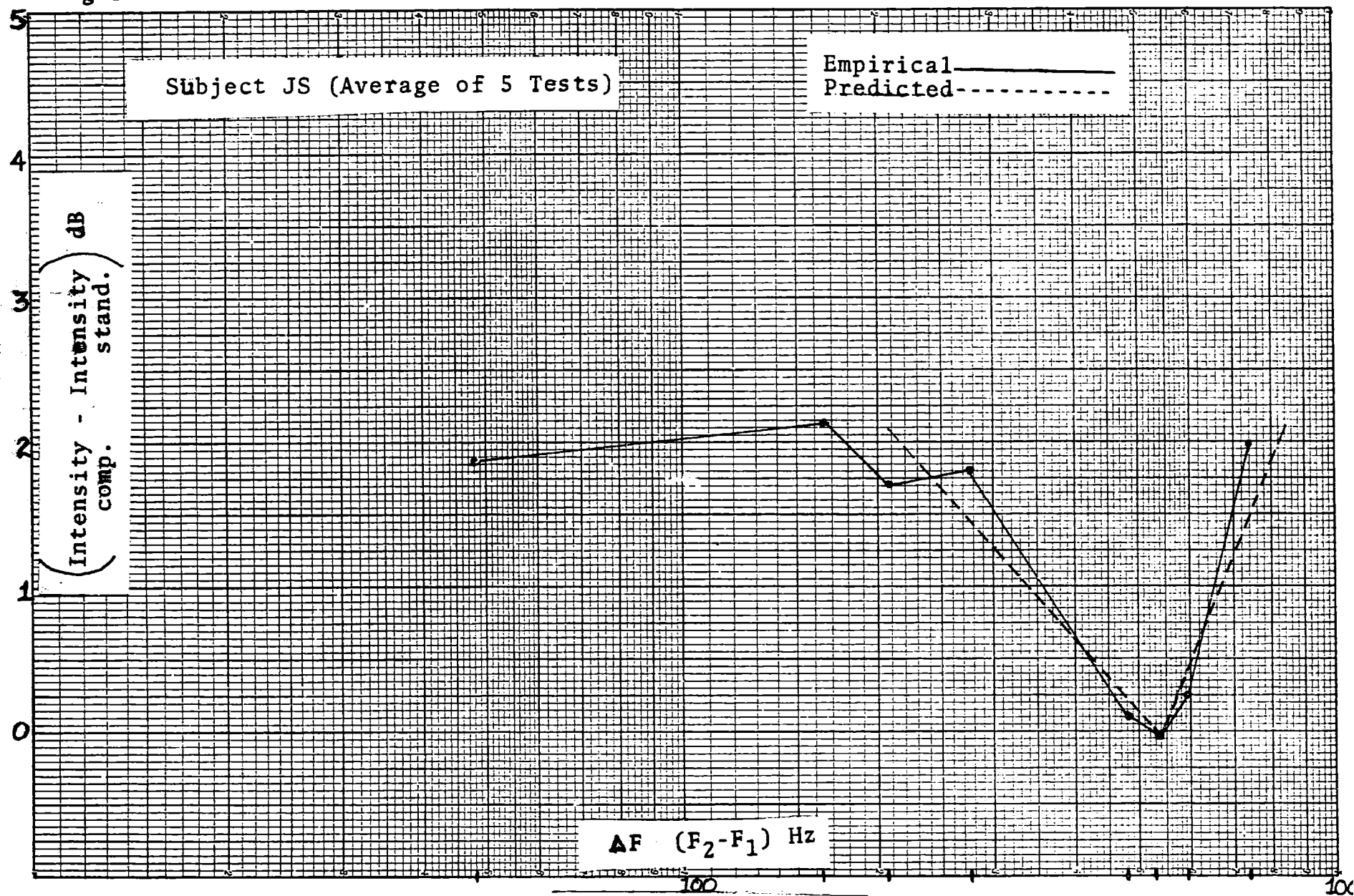


100

10

Figure #15.

DICH-CONS



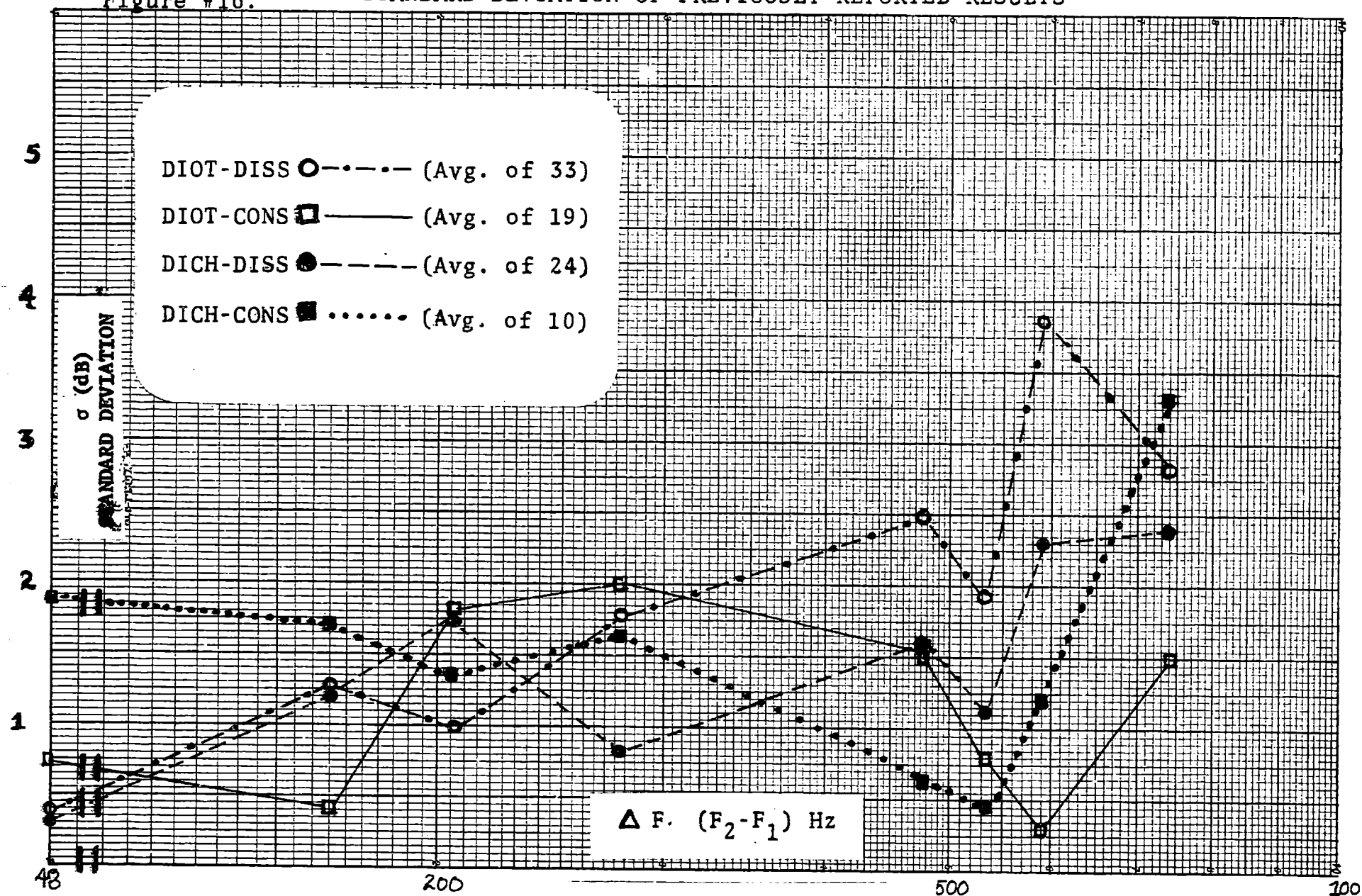
the loudness determinations for each ΔF tested are shown on Figure #16 and they certainly seem to agree with the above hypothesis. The similarity between these curves and their respective equiloudness curves presented (Figures #4-#16) is almost uncanny. The fact that the shape of the dich-cons case for both types of curves has almost exactly the same trends, with the curve being high at the smaller ΔF 's and showing a somewhat large depression centered at the ΔF of a 2:3 ratio, indicates that the 'ambiguity' of the sound is directly proportional to the loudness estimation. Other interesting similarities exist for the other three cases and their respective curves; since a very definite 'dip' occurs at the consonant and for small

ΔF 's, the values of the intensity - intensity are low.
comp. stand.

Therefore, with the above explanation, there isn't necessarily any inconsistency with this last case (dich-cons) concerning the disappearance of the 'dip' at the consonant. Harmonically-oriented sounds have been known²¹ to have much smaller pitch ambiguities (or pitch confusions) than comparable sounds and this of course, is in accordance with this hypothesis. On the other hand, generally speaking, dissonant sounds have much greater pitch ambiguities (as can be seen analytically from Goldstein's pitch model²² and experimentally from the work of Smoorenberg²³ and Plomp²⁴) and this also is very much in agreement with this hypothesis as well as the observations. Almost everything seems to fall into place. Since the loudness ambiguity resembles the average loudness estimate quite closely, and if we assume that the loudness ambiguity is directly related

Figure #16.

STANDARD DEVIATION OF PREVIOUSLY REPORTED RESULTS



to the pitch ambiguity, the loudness estimate is therefore some function of the pitch confusion also. This assumption does not seem too far fetched since it was determined here that the loudness of two-tone complexes is some function of pitch.

The diotically presented stimuli produce equiloudness curves that in general contain higher values of the intensity - intensity ^{comp.} probably because of the increase in confusion _{stand.} due to the audible nonlinear tones present. There seems to be two phenomena that add in some way when judging loudness. The first is the general trend of the **change** in loudness with changing pitch (very similar to that of the **Monotone** effect), and the second is the increase in loudness with increasing pitch confusion or ambiguity. This latter idea is of course, not directly shown here, but with each ΔF , there is a corresponding pitch and confusion, and if taken in that light, the path of the aforementioned hypothesis is indeed plausible.

It should be noted here that the predicted curves are the author's estimation of the equiloudness curves with the above hypothesis as the constraints. Although the pitch determination has not been made for each two-tone complex, and the confusion matrix has not yet been formed for these sounds, the predicted values were obtained from a very gross approximation of these.

Chapter 5

CONCLUSION AND DIRECTIONS FOR FUTURE WORK

It has been shown here that a harmonic two-tone complex (musically defined) appears softer to the human perception than a comparable dissonant sound (close in pitch with higher ambiguity). This idea, although shown qualitatively has not been quantitatively done in any way. This has been left to future work.

The methods to support the hypothesis that loudness is indeed a function of intensity, pitch, and pitch ambiguity should use the fact ~~that~~ harmonically-oriented sounds are almost by definition a lot less ambiguous with respect to pitch than a dissonantly-oriented sound. This seems to agree with the observed fact that the loudness of music and jet engine noise at the same intensity are judged differently. It should be obvious that the music from the symphony orchestra has a much more defined pitch structure than that of the jet engine.

As stated previously, pitch formation has been found to be central. Since pitch ambiguity must also originate centrally, the fact that loudness (at least for two-tone complexes) is a result of some central phenomenon should not be too surprising.

The idea that loudness is some function of pitch is not a new one. The idea that annoyance, or unpleasantness, or "lacking musical quality", is directly related to pitch ambiguity is new. The idea that an increase in loudness is related

in some complicated way to an increase in unpleasantness can be explained by the fact that pitch ambiguity becomes larger with increasing intensity (as seen before because of aural harmonics). The question of which comes first, the loudness ambiguity or the pitch ambiguity, does not seem to be too important for us to answer. What does seem very important is if we know one, can we get the other? If we know and can understand the reasons for one, does this give us an important clue to finding the cause for the other? If we know the intensity, pitch, and pitch ambiguity of a complex sound, can we predict with a much greater accuracy the loudness? If we know the frequencies, loudness, and loudness ambiguity, can we predict the pitch? From the foundation laid here, with more experimentation, the author believes we can.

If we could quantify a sound as to its musicality or noisiness (pitch ambiguity), then we can communicate and specify a complex sound as to its annoyance or loudness more precisely. With the advent of some kind of loudness-pitch theory, this end can be accomplished.

If it is found that a complex sound can be made 'quieter' through some sort of pitch or pitch ambiguity change, instead of an overall power cutback (lower intensity), then there are amazing possibilities. Imagine a jet engine sounding consonant; it is bound to be at least tolerable. Then maybe, with greater technology, it can have a controlled timing structure and intensity variation. A musical jet! Why not? After all, the

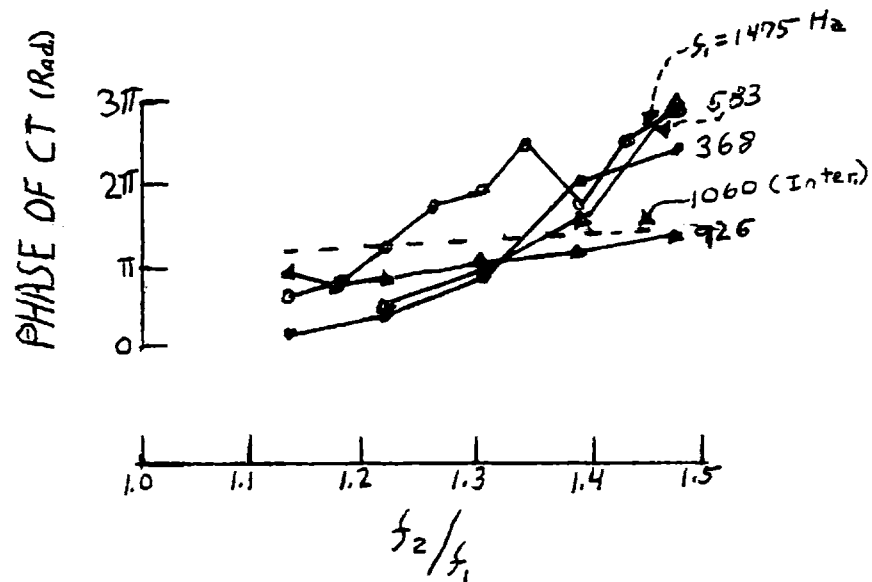
world could certainly do with alot less noise and alot more music!

APPENDIX A

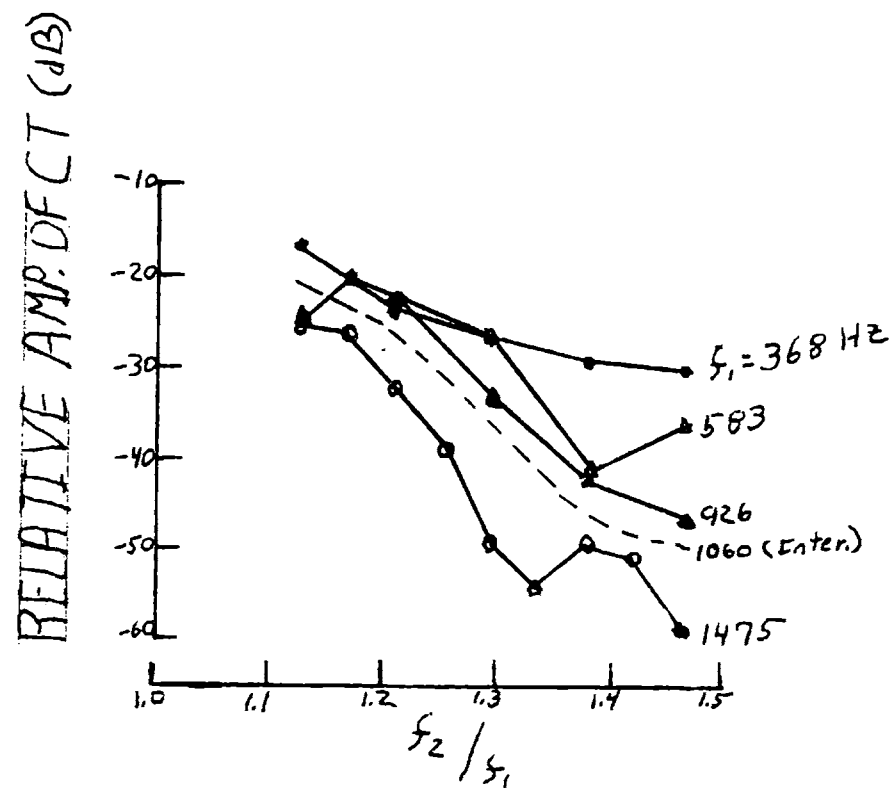
As can be seen quite clearly from the figures from J. L. Hall's paper, presented on the following page, for the intensity used here (68 dB S.P.L.) and the frequency ratio of 2:3 or 1.5, the difference tone ($f_2 - f_1$) is approximately 55 dB down from the primaries and the combination tone ($2f_1 - f_2$) is approximately 50 dB down. Although these two tones would be barely audible, the interference between them (where the difference tone has a phase of about 180° with respect to the primaries and the combination tone has a phase of about -180°) makes their added intensity even lower.

It can also be seen from these graphs that at the minor third (5:6), the combination tone is as high as 30 dB down from the primaries, and it is here that they would be a significant factor in pitch or loudness perception.

In practice, however, for real musical instruments, most musical intervals contain very audible nonlinear tones due to the many overtones present per note. In this paper, when we refer to 'musical interval', we are speaking of two-tone complexes only.

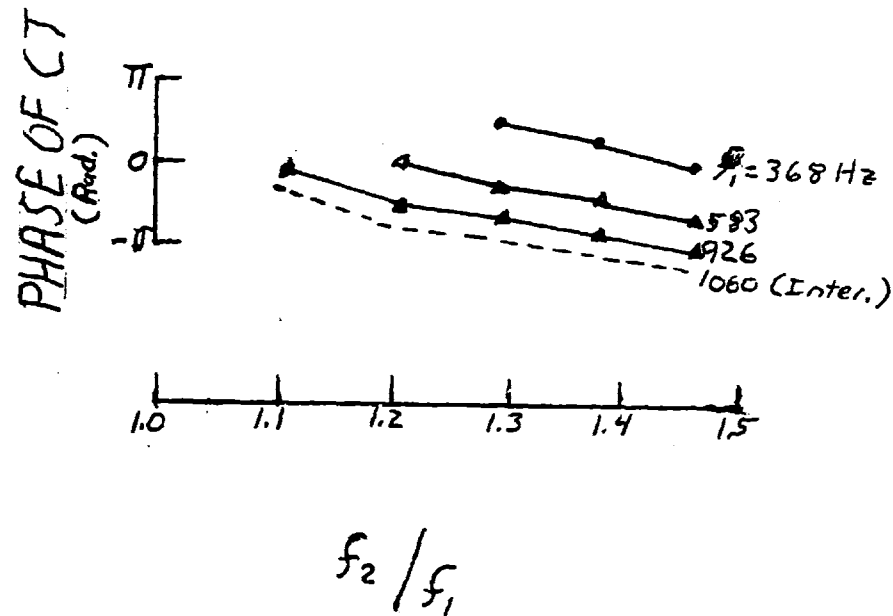


(a)



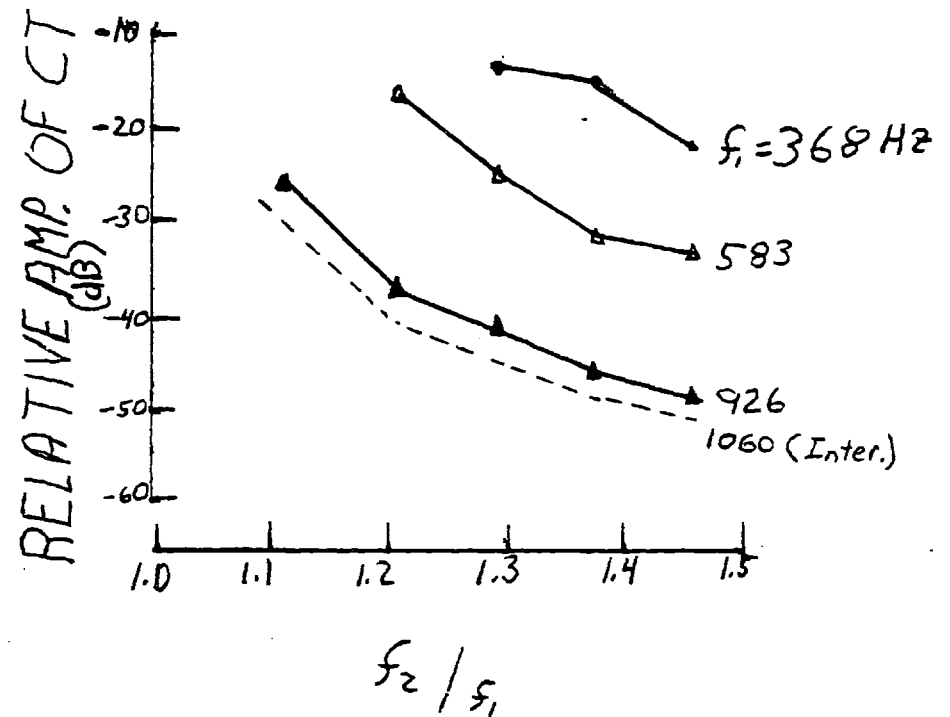
(b)

Figure A. Amplitude and phase of $2f_1 - f_2$ CT, with primary tones at 68 dB S.P.L. The phase angle of the CT is relative to the primary tones and shown in (a) and the amplitude of the CT relative to the primary tones is shown in (b).



f_2/f_1

(a)



f_2/f_1

(b)

Figure B. Amplitude and phase of f_2-f_1 CT, with primary tones at 68 dB S.P.L. The phase angle of the CT is relative to the primary tones and shown in (a) and the amplitude of the CT relative to the primary tones is shown in (b).

REFERENCES

1. Webster's Seventh New Collegiate Dictionary, G. & C. Merriam Company, Mass., 1967, 572.
2. Plomp, Smoorenburg, eds. Frequency Analysis and Periodicity Detection in Hearing, A. W. Sijthoff Leiden, 1970.
3. Houtsma, A. J. M., Goldstein, J. L., "The Central Origin of the Pitch of Complex Tones: Evidence from Musical Interval Recognition", J.A.S.A. 51, 1971, 520-529.
4. Goldstein, J.L., "An optimum processor theory for the central formation of the pitch of complex tones", J.A.S.A. 54, 1973, 1496-1516.
5. Plomp, "Pitch of Complex Tones", J.A.S.A. 41, 1967, 1526-1533.
6. See #3.
7. a) Schouten, J.F., "The Perception of Pitch", Philips Tech. Rev. 5, 1940, 286-294.
b) Schouten, J.F., Ritsma, R.J. and Cardozo, B.L., "Pitch of the Residue", J.A.S.A. 34, 1962, 1418-1424.
8. Kryter, Karl D., The Effects of Noise on Man, New York: Academic Press, 1970.
9. a) Stevens, S.S., Handbook of Experimental Psychology, Wiley & Sons, N.Y., 1951.
b) Stevens, S.S., "The Relation of Pitch to Intensity", J.A.S.A. 6, 1939, 150-159.

10. Goldstein, J.L., "Auditory Nonlinearity", J.A.S.A. 41, 1967, 679-689.
11. Hall, J.L., "Monaural Phase Effect: Cancellation and Reinforcement of Distortion Products f_2-f_1 and $2f_1-f_2$ ", J.A.S.A. 51, 1972, 1872-1881.
12. a) Hall, J.L., "Auditory Distortion Products f_2-f_1 and $2f_1-f_2$ ", J.A.S.A. 51, 1972, 1863-1871.
b) See Appendix A.
13. Zwicker, E., Flottorp, G., and Stevens, S.S., "Critical bandwidth in loudness summation", J.A.S.A. 29, 1957, 548-557.
14. See #2.
15. Plomp suggested that frequencies separated by more than a critical band become more pleasing to listen to. Since combination tones decrease in intensity as ΔF increases, this does not seem at all contradictory.
16. See #13.
17. The fact that we are assuming that sounds containing frequencies related by exact integer ratios are identical to musical consonances is not exactly true. Since in reality most musical instruments contain many overtones as well as the advent of the tempered scale, a slight mistuning of frequencies results in producing the consonant chords.
18. Plomp, R., "Beats of Mistuned Consonances", J.A.S.A. 42, 1967, 462-474.

19. It should be noted here that additional experiments resembling very closely those described here (they used a 1000 Hz tone as the standard, and their F_2 's were slightly different with a $F_1 = 1000$ Hz instead of 1060 Hz) were done under the direction of J.K. Haviland, Dept. of Engr. Science and Systems, University of Virginia in the spring of 1974. For thirteen subjects, all but one curve significantly showed this 'dip' described here.
20. See #'s 10, 11, 12.
21. Smoorenburg, G.F., "Pitch Perception of Two-Frequency Stimuli", J.A.S.A. 48, 1970, 924-941.
22. See #4.
23. See #21.
24. See #5.